

# ***TeraGrid***

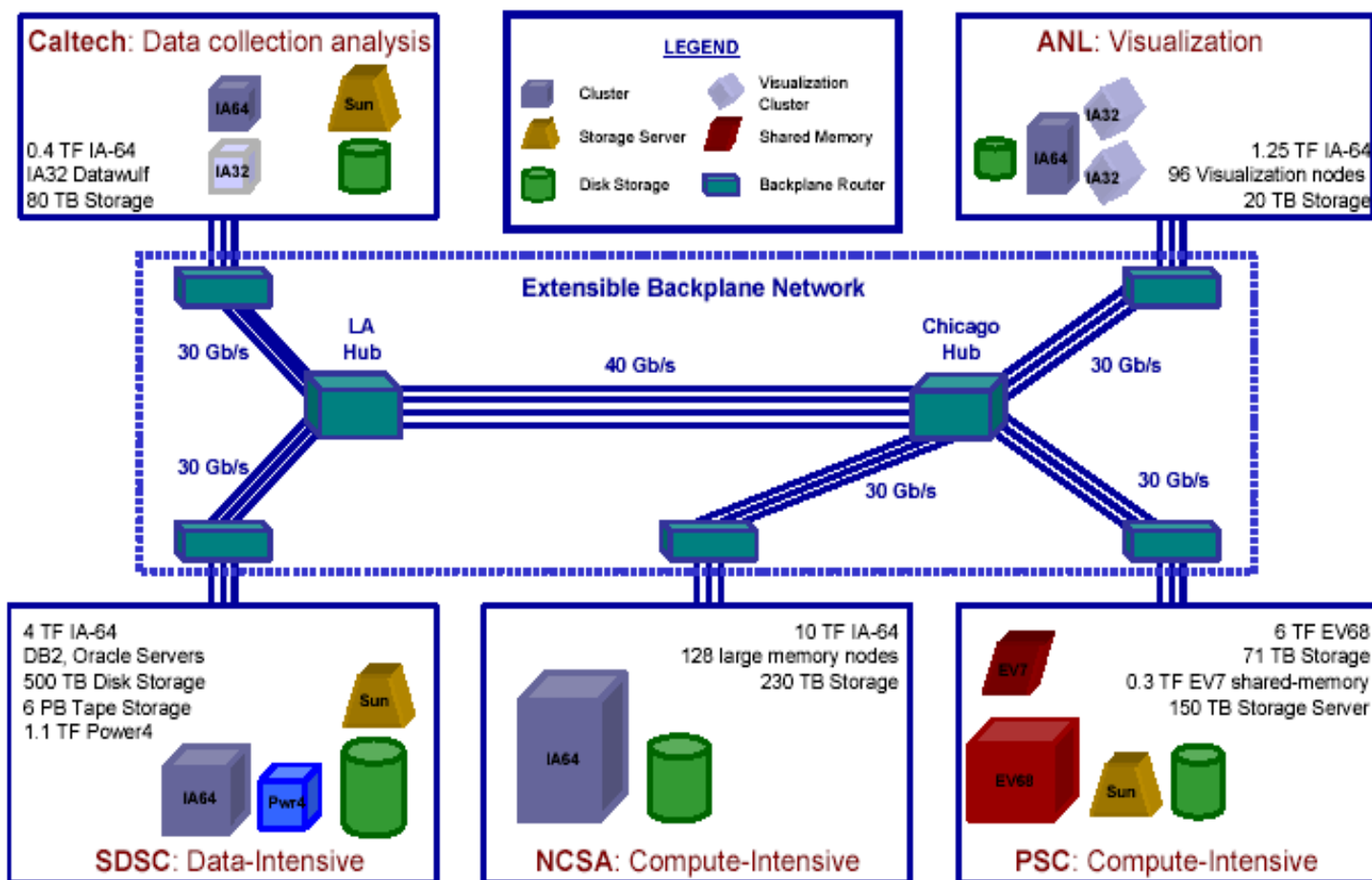
**a powerful new research tool for  
digital biology**

**Richard L. Hilderbrandt  
PACI Program Director**





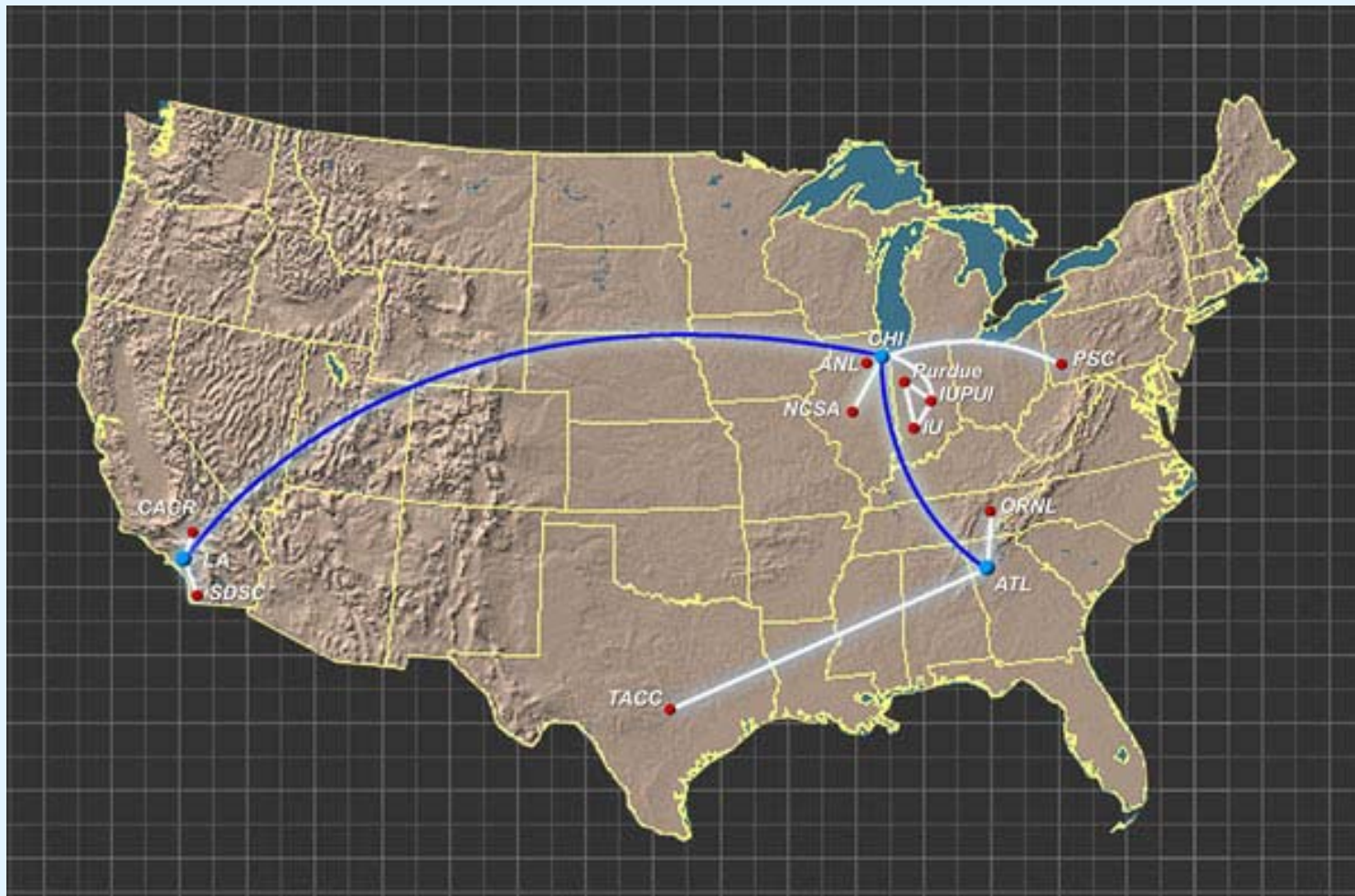
# TeraGrid (ETF) Configuration





# *The NSF Extensible Terascale Facility*

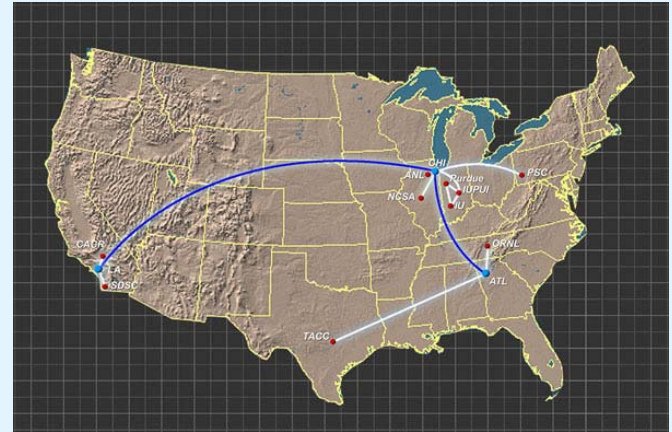
---





## *ETF Objectives*

- Create unprecedented capability
  - integrated with extant PACI capabilities
  - supporting a new class of scientific research
- Deploy a balanced, distributed system
  - not a “distributed computer” but rather
  - a distributed “system” using Grid technologies
    - computing and data management
    - visualization and scientific application analysis
- Define an open and extensible infrastructure
  - an “enabling cyberinfrastructure” for scientific research
  - extensible beyond original sites with additional funding
    - NCSA, SDSC, ANL, Caltech and PSC
    - other sites via expanded ETF solicitation





## *Middleware Enables*

---

- Integration of remote resources and collaboration capabilities into local experimental, computational, and visualization environments
- Discovery and utilization of scientific data, computers, software and instruments over the network in a controlled manner
- Averting or mitigating failures or poor performance in distributed systems
- Managing the authoring, publication, curation and evolution of scientific data, programs, computations and other products.



## *Science is enhanced by:*

---

- Computing capacity adequate for a task at the time the task is needed by the science
- Data capacity sufficient for the science task independent of location.
- Remote access to distributed computing, data, instrumentation, and collaboration tools for a distributed science community
- Software services that give scientists the ability to build simulations that are natural to the scientific process rather than having to focus on the underlying computing, data and communication resources
- Communication capacity and capability sufficient to support all of this transparently to both systems and users.



## *Focus on Policy and Social Dynamics*

---

- Policy issues must be considered up front
- Social engineering will be at least as important as software engineering
- Well-defined interfaces will be critical for successful software development
- Application communities will need to participate from the beginning

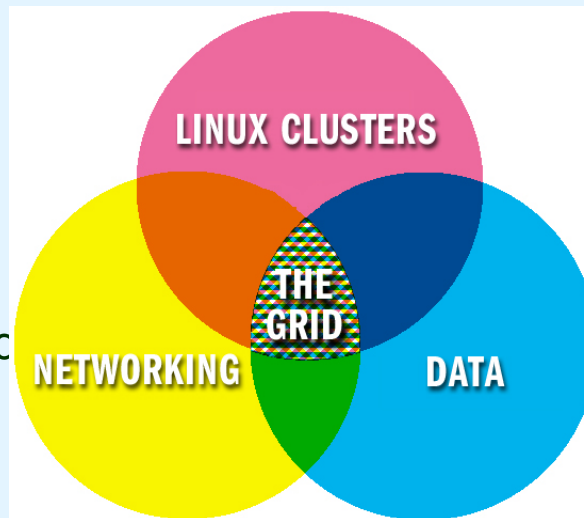






# *ETF Common Infrastructure Environment*

- Linux Operating Environment
- Basic and Core Globus Services
  - GSI (Grid Security Infrastructure)
  - GSI-enabled SSH and GSIFTP
  - GRAM (Grid Resource Allocation & Management)
  - GridFTP
  - Information Service
  - Distributed accounting
  - MPICH-G2
  - Science Portals

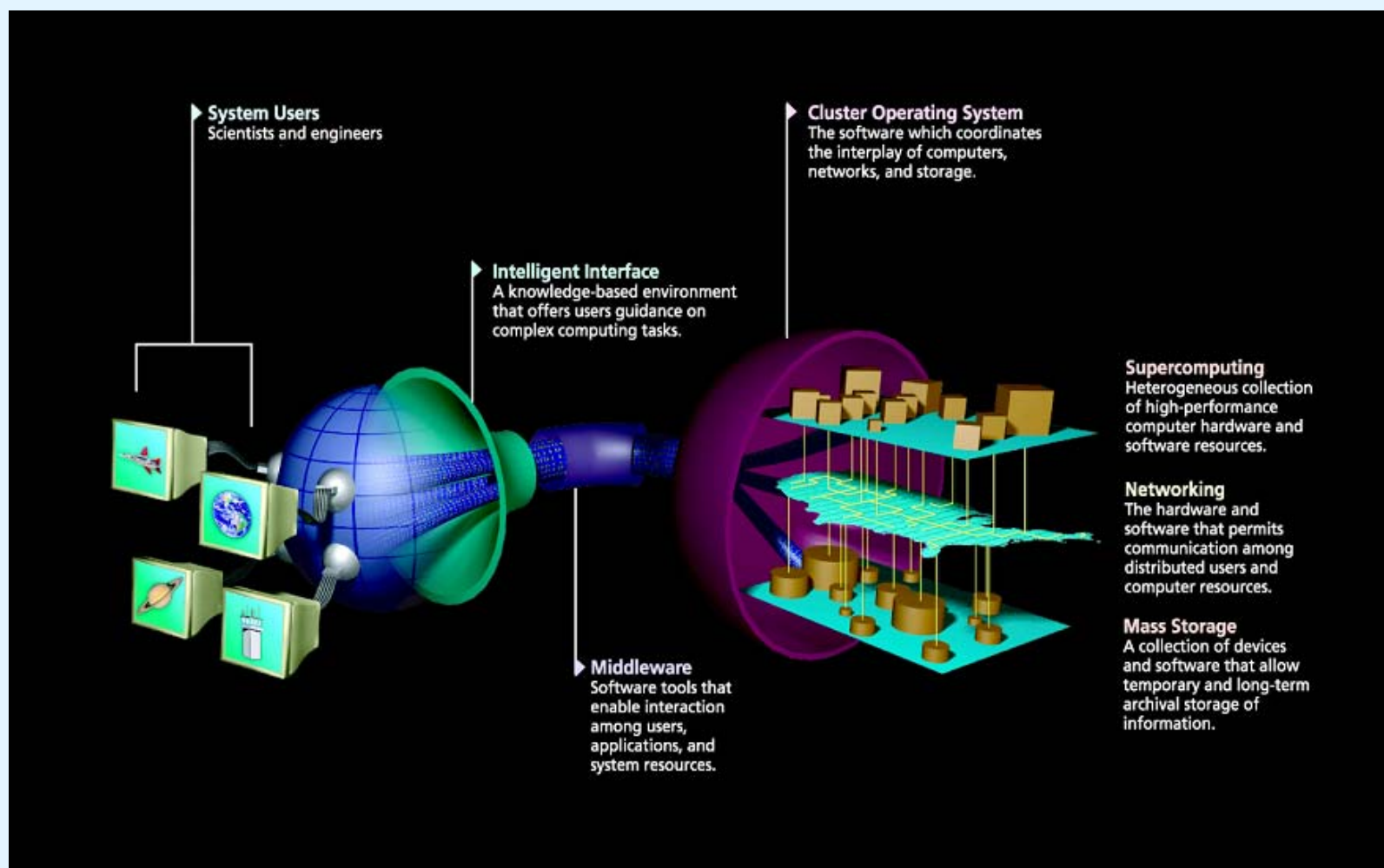


- Advanced and Data Services
  - Replica Management Tools
  - GRAM-2 (GRAM extensions)
  - CAS (Community Authorization Service)
  - Condor-G (as brokering “super scheduler”)
  - SDSC SRB (Storage Resource Broker)
  - APST user middleware, etc.





# *Vision of Grid Computing*



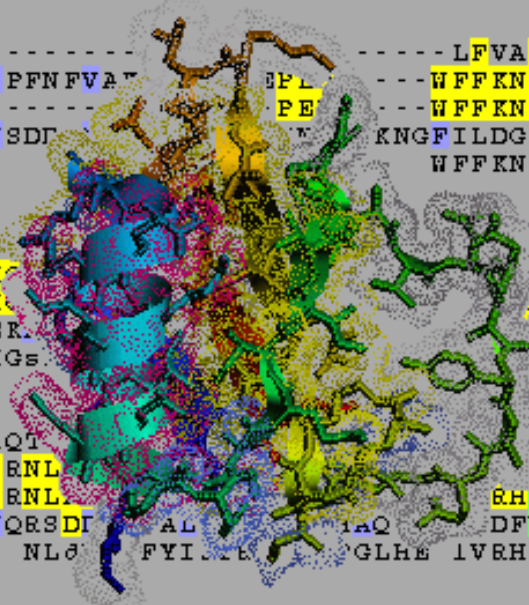


# Biology WorkBench

click here to  
toggle between  
menus and buttons

**WE  
Moved!** <http://workbench.sdsc.edu/>

Version 3.2



```
-----LFVALY-----  
QSLTTGQEGFIPFNFA-----WFFKNLSRK--DAER--  
-----WFFKNLSRK--DAER--  
AKKIMDQGGGLVSD-----KNGFILDGFPTIPQAEKLD  
      Q G I      WFFKNLSRK  DAER  
  
DFVASGD-----YNHN-----  
QLLAPGNT--HG--FDQN-----  
QLLAPGNT--HG--FDQN-----  
QMLKEQGTPLEK-----YHKIFNPPKED  
QLIAPGNT  HG--yDQN-----  
  
--GE--WCEAQT-----IT-----  
--QGEVVVKHYKIRNL-----TN-----  
--QGEVVVKHYKIRNL-----SHYTN-----  
MKDDVTGEALVQRSDFAL--IAQ--DFYKKTGIWAGVDA  
QGEVVVKHYKIRNL--FYIS--PGLHE IVERH TN
```

Session Tools

Protein Tools

Nucleic Tools

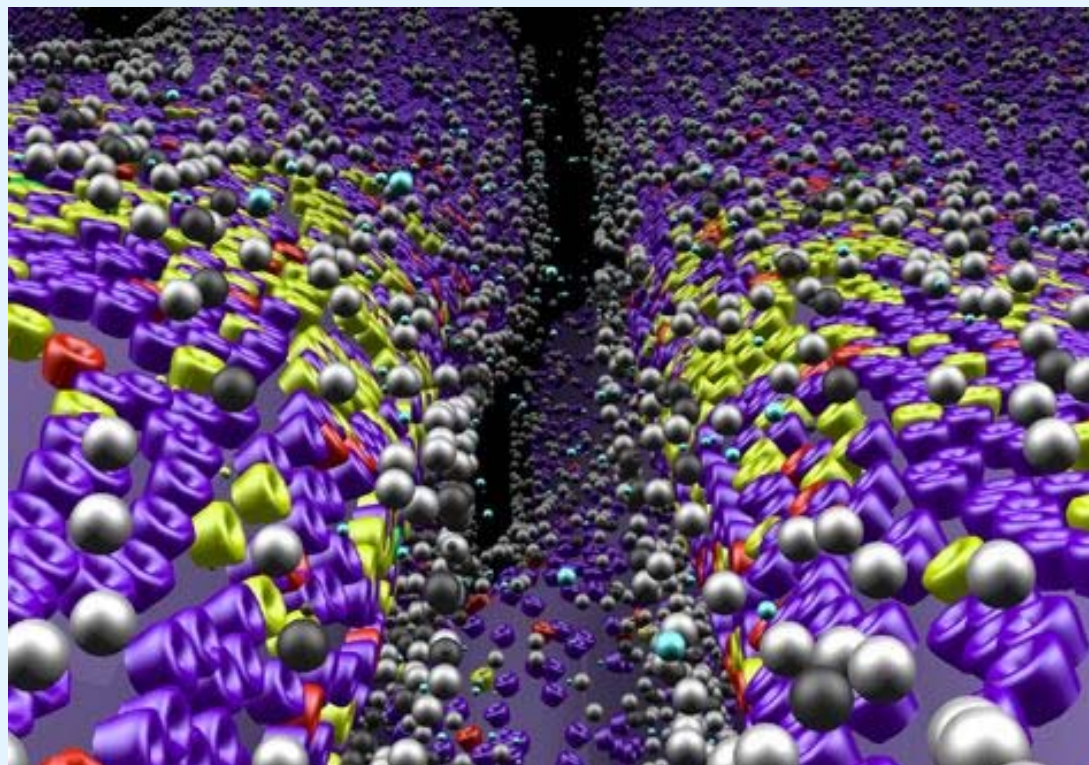
Alignment Tools

Structure Tools (Alpha)

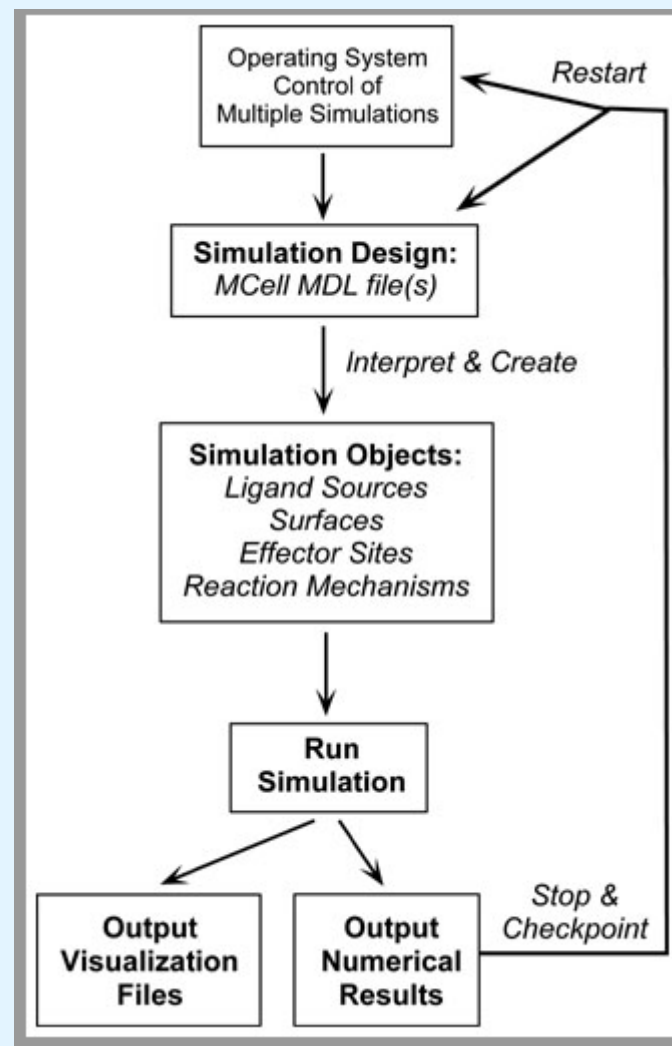
Default Session



# MCELL: Monte Carlo Simulator of Cellular Microphysiology



Tom Bartol and Joel Stiles

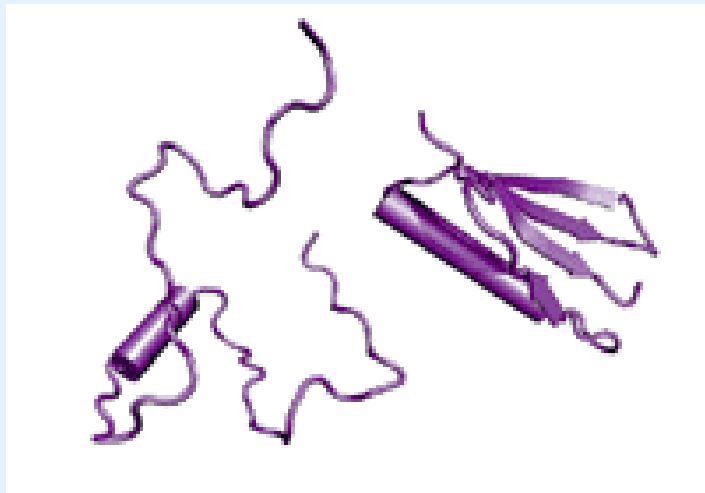




# Protein Folding in a Distributed Computing Environment

---

Project Leaders: Marty Humphrey, *University of Virginia*; Charles L. Brooks III, *The Scripps Research Institute*  
Project Manager: Marty Humphrey, *University of Virginia*



## Biochemical Origami

Protein L unfolded (left), and folded (right) in a CHARMM simulation.

A Web portal enables any AMBER user to sign on to NPACI-net. Input files can come from local directories or Legion's global distributed file system, and runs can be monitored, steered, and interrogated to produce intermediate files. The portal also enables visualization of results.





# telescience

for advanced tomography applications

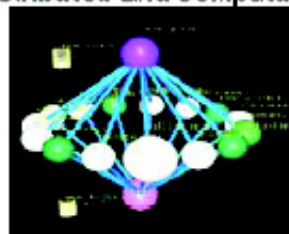
## telescience

for advanced tomography applications

Remotely Control Instruments



Distributed Grid Computation



Distribute Data to Storage Resources



Username

Password

Login

[What is Telescience?](#)

[Learn more about the Telescience Portal](#)

[Need a Telescience Account?](#)



NPACI





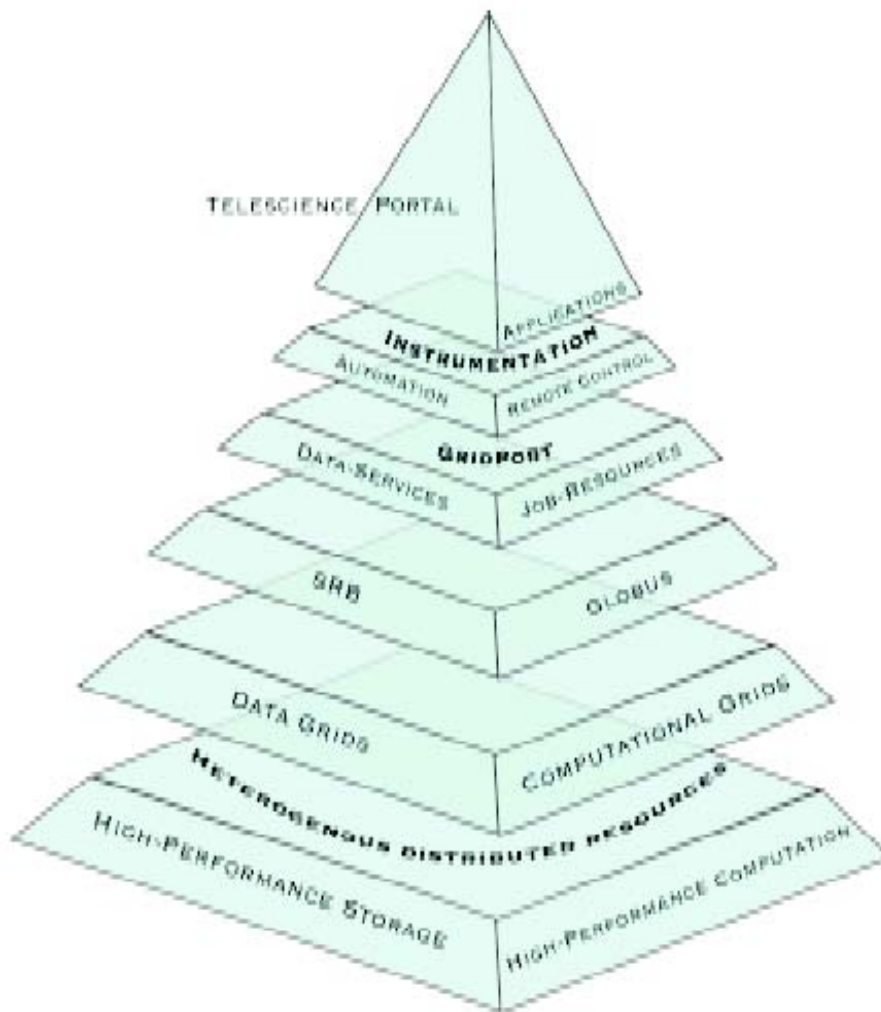
# *Telescience Portal Capabilities*



- experiment planning and organization,
- equipment and resource scheduling,
- Telemicroscopy,
- image processing,
- parallel tomographic volume reconstruction,
- visualization and analysis,
- data management,
- and database population.



# Telescope Portal Layered Structure



- Telescope Portal is composed of many “layers”
- Layers are modular, allowing for extension of each layer without disrupting the entire system
- Every Layer has its own complexity and administration that was previously passed on to the end-user
- Telescope Portal centralizes all administrative details of each layer into a single username and passphrase





## *BioCoRE Job Submission Features*

---

- **Job Manager** that allows scientists to submit jobs to National Science Foundation high-performance computers, including those at NCSA.
- An **interface to NAMD**. BioCoRE makes it easier to create and edit NAMD input files.
- Users can begin viewing the output of their computations while the job is still running, which can lead to early detection of flawed input and can save precious machine hours.



## *BioCoRE Collaboratory Features*

---

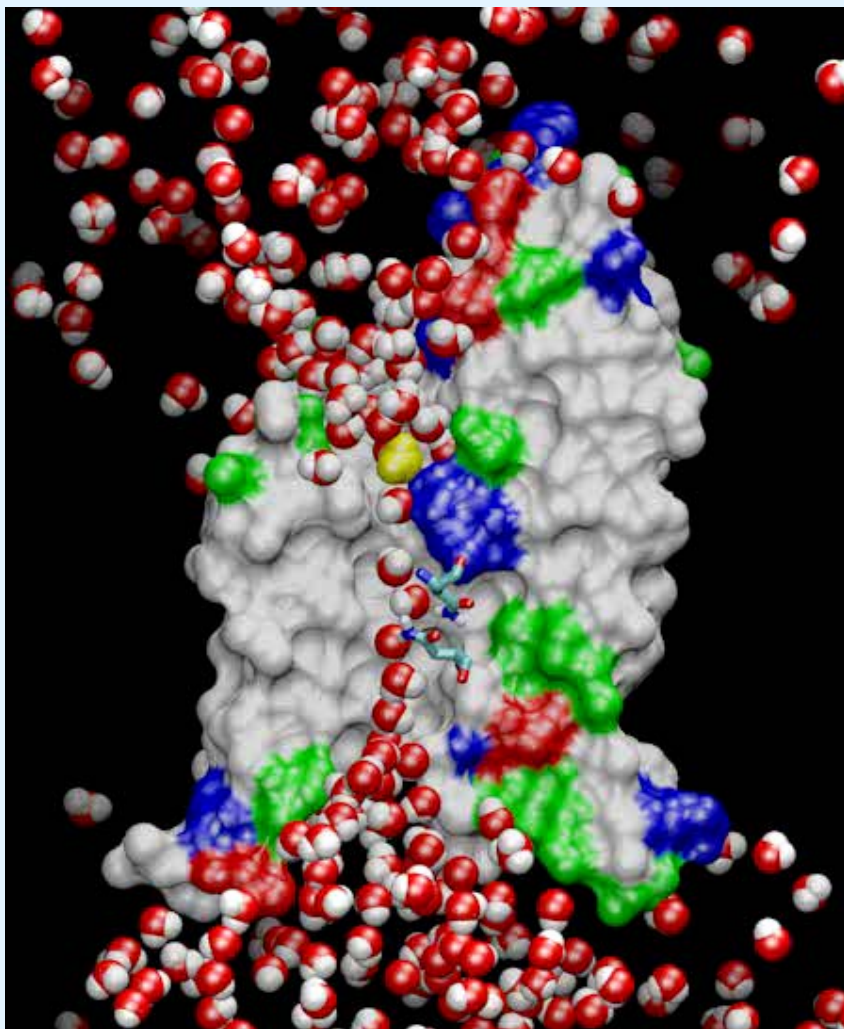
- Tools to save and share the **3D molecular visualizations** created in VMD.
- A **shared file system**. Shared files are represented by the same type of icons as files and programs that live on a user's desktop computer.
- A **message board** and online **lab notebook** to help distant collaborators communicate and stay in sync.



## *2003 Nobel Prize in Chemistry*

### *Peter Agre Aquaporins*

---



### Aquaporin Simulation by Klaus Schulten

Drs. Emad Tajkhorshid and Klaus Schulten. Tajkhorshid, E., Nollert, P., Jensen, M.O., Miercke, L.J., O'Connell, J., Stroud, R.M., and Schulten, K. (2002). *Science* 296, 525-530.

Simulation performed at  
using the Terascale  
Computing System at  
PSC.



Grid Research Integration Deployment and Support Center

**GRIDS**



Define, **Develop,**  
Deploy, and  
Support an  
**Integrated**  
national middleware  
infrastructure  
supporting  
**21st Century**  
science and  
engineering  
**Applications**

- [\*\*NMI-R3.2\*\*](#) is the latest release from the NSF Middleware Initiative (NMI), featuring the [\*\*GRIDS Center Software Suite\*\*](#) and contributions from the NMI-EDIT team. New GRIDS components include MyProxy, GridConfig, and MPICH-G2.



# *NMI-EDIT*

## *Enterprise and Desktop Integration Technologies*

---



- Foster interoperability among security and privacy management architectures
- Encourage a coherent middleware architecture design that integrates with individual desktops and operating systems.
- Enable new valuable applications for research and higher education.
- Educate to insure the use of middleware services across higher education, vendor, and research communities.